Structural and functional diversity of substrate-binding domains in ABC importers

Fulyani, Faizah

Chapter 2
Functional diversity of tandem substrate-binding domains in ABC transporters from pathogenic bacteria

Faizah Fulyani, Gea K. Schuurman-Wolters*, Andreja Vujičić Žagar*, Albert Guskov, Dirk-Jan Slotboom and Bert Poolman*

This chapter has been published in:
Structure (2013) **21**:1879-1888

* these authors contributed equally to this work

Abstract

The ATP-binding cassette (ABC) transporter GlnPQ is an essential uptake system for amino acids in Gram-positive pathogens and related non-pathogenic bacteria. The transporter has tandem substrate-binding domains (SBDs) fused to each transmembrane domain, giving rise to four SBDs per functional transporter complex. We have determined the crystal structures and ligand binding properties of the SBDs of GlnPQ from Enterococcus faecalis, Streptococcus pneumoniae and Lactococcus lactis. The tandem SBDs differ in substrate specificity and affinity, allowing cells to efficiently accumulate different amino acids via a single ABC transporter. The combined structural, functional and thermodynamic analysis revealed the roles of individual residues in determining the substrate affinity. We succeeded in converting a low-affinity SBD into a high-affinity receptor and vice versa. Our data indicate that a small number of residues that reside in the binding pocket constitute the major affinity determinants of the SBDs.

1.1 INTRODUCTION

ATP-binding cassette (ABC) transporters have been subdivided into exporters and Type I, II and III importers based on the structures of their integral membrane domains [1,2]. Exporters mediate transport of molecules from the cytoplasm to the external medium or organelle lumen. They bind their ligands directly within the transmembrane domain (TMD), without the need for auxiliary proteins. Type I and II importers capture their ligands via soluble substrate-binding proteins (SBP). Type III importers, also known as ECF-type ABC transporters [3], capture a substrate via membrane-embedded S-components [4], associated with the energy-coupling factor (ECF). Upon substrate binding, the SBPs of Type I and II importers change conformation from open to closed and, subsequently, dock onto the TMD. Generally, both Type I and II importers follow a so-called "two state" alternating access model in which the transporter adopts either an outward- or inward-facing conformation, thus allowing substrate to be transferred from the external medium to the cytoplasm. The mechanism of transport of Type I and II importers is quite different. In Type I, as exemplified by the maltose transporter MalE-MalFGK$_2$ from *Escherichia coli* [5,6], liganded MalE interacts with low affinity to inward-facing MalFG, which then triggers ATP binding and closure of the nucleotide-binding domains (NBDs) and the transfer of substrate from the SBP to the TMD [7,8]. Subsequent hydrolysis of ATP and release of inorganic phosphate (Pi) and ADP completes the translocation cycle and resets the system to the ground-state. In an alternative model, binding of ATP triggers the outward-facing conformation of MalFG to which unliganded MalE binds with high affinity [9]. Binding of maltose to MalE-MalFGK$_2$ would then initiate substrate translocation. In Type II importers (*e.g.* the vitamin B12 transporter, BtuF-BtuCD), the transport is initiated by docking of liganded ButF to the outward-facing conformation of the transporter [10–12]. Binding of ATP closes the NBDs as well as the periplasmic and cytoplasmic gates, and the substrate gets trapped in a translocation cavity ("occluded state"). Subsequent ATP hydrolysis opens the cytoplasmic gate and releases the substrate on the *trans* side of the membrane [10].

Soluble SBPs were first discovered in the periplasm of the Gram-negative bacterium *E. coli*, thus they are often referred to as periplasmic binding proteins [13]. In microorganisms lacking an outer membrane and periplasm, *i.e.* Gram-positive bacteria and Archaea, SBPs are exposed on the cell surface. Typically they are attached to the cytoplasmic membrane via either a lipid-anchor or a membrane-embedded peptide (the latter has only been observed in Archaea), or they are fused to the TMDs resulting in two substrate-binding domains (SBDs) per

functional complex. In some cases, two or even three SBDs are fused together and linked to the TMDs generating in total four or six extracytoplasmic substrate-binding sites. Transporters with SBDs fused to the TMDs occasionally are also present in Gram-negative bacteria but less frequently than in Gram-positives [14]. The linkage of SBDs to the membrane and fusing of (multiple) SBDs to the TMD increase the effective concentration of the substrate-binding sites near the translocator. SBDs are not only associated with ABC transporters but also present in ion-linked transporters, ion channels, G-protein coupled receptors and two-component regulatory systems, and they are found in prokaryotes and eukaryotes [15]. Despite their importance in biology and the availability of many crystal structures, the mechanism of ligand binding is still poorly understood, thus it is difficult to predict ligand specificity from protein structure and therefore to rationally design drugs. What is needed is a better description of the functional and thermodynamic properties of SBDs in relation to structural information.

Here, we focus on the functional and structural analysis of Type I ABC importers with two SBDs fused in tandem (termed as SBD1 and SBD2 for the individual domains and SBD12 for the tandem), which are prominent in Gram-positive pathogenic bacteria. These transporters are among the most complex systems of the ABC superfamily, and the functional role of dual SBDs is poorly understood. We reveal the specificity determinants of the SBDs by analysis of the orthologous GlnPQ importers from the non-pathogenic *L. lactis* and the pathogens *Enterococcus faecalis* and *Streptococcus pneumoniae.* We find that the substrate specificity of the two SBDs in the tandem is different, with one SBD preferring asparagine and the other glutamine. Both SBDs can bind glutamine but the affinities differ by four orders of magnitude, even though the amino acid sequence and tertiary structures are highly conserved. Using a combined structural biology and thermodynamic approach and structure-based engineering, we identified the determinants for substrate affinity of the tandem SBDs of the GlnPQ importers from *L. lactis*, *E. faecalis* and *S. pneumoniae*.

1.2 RESULTS

1.2.1 Gene organization, cloning, and protein purification

The GlnPQ importers from *L. lactis* (Ll), *E. faecalis* (Ef) and *S. pneumoniae* (Spn) are composed of two subunits: GlnP and GlnQ, and each subunit is utilized twice in the functional transporter complex. GlnP consists of a signal sequence (SQ), two SBDs fused in tandem and a C-terminal TMD. GlnQ is a typical ABC-type nucleotide-binding protein. Figure 1A shows the gene organization of the *glnPQ* operon and

the neighboring genes in *L. lactis*, *E. faecalis* and *S. pneumoniae*. Gene *glnP* is followed immediately by *glnQ* and both are expressed from *the same promoter.* The sequence identity among tandem SBDs of the GlnP proteins from *L. lactis*, *E. faecalis* and *S. pneumoniae* is about 50 % with major differences in the linker region that connects SBD1 and SBD2 (Figure 1B). The residues that interact with the substrate in the crystal structures of SBD1 and SBD2 from *L. lactis* are highlighted in red and are actually well conserved in the other proteins (*vide infra*).

The genes corresponding to the individual SBDs and the tandem SBD from *E. faecalis, S. pneumoniae* and *L. lactis* were expressed in *E. coli* [16]. In each case, a N-terminal decahistidine tag followed by the TEV cleavage site preceded the sequence of the mature protein. A two-step purification procedure (Immobilized Metal Affinity Chromatography followed by Size-Exclusion Chromatography) typically yielded more than 10 mg of monodisperse protein from 1 liter of cell culture ($OD_{600} \sim 2.8$), both for the individual and tandem SBDs. The purified SBDs were monomeric with a molecular mass of ~ 25 kDa; the tandem SBDs also behaved as monomeric species (Figure 2).

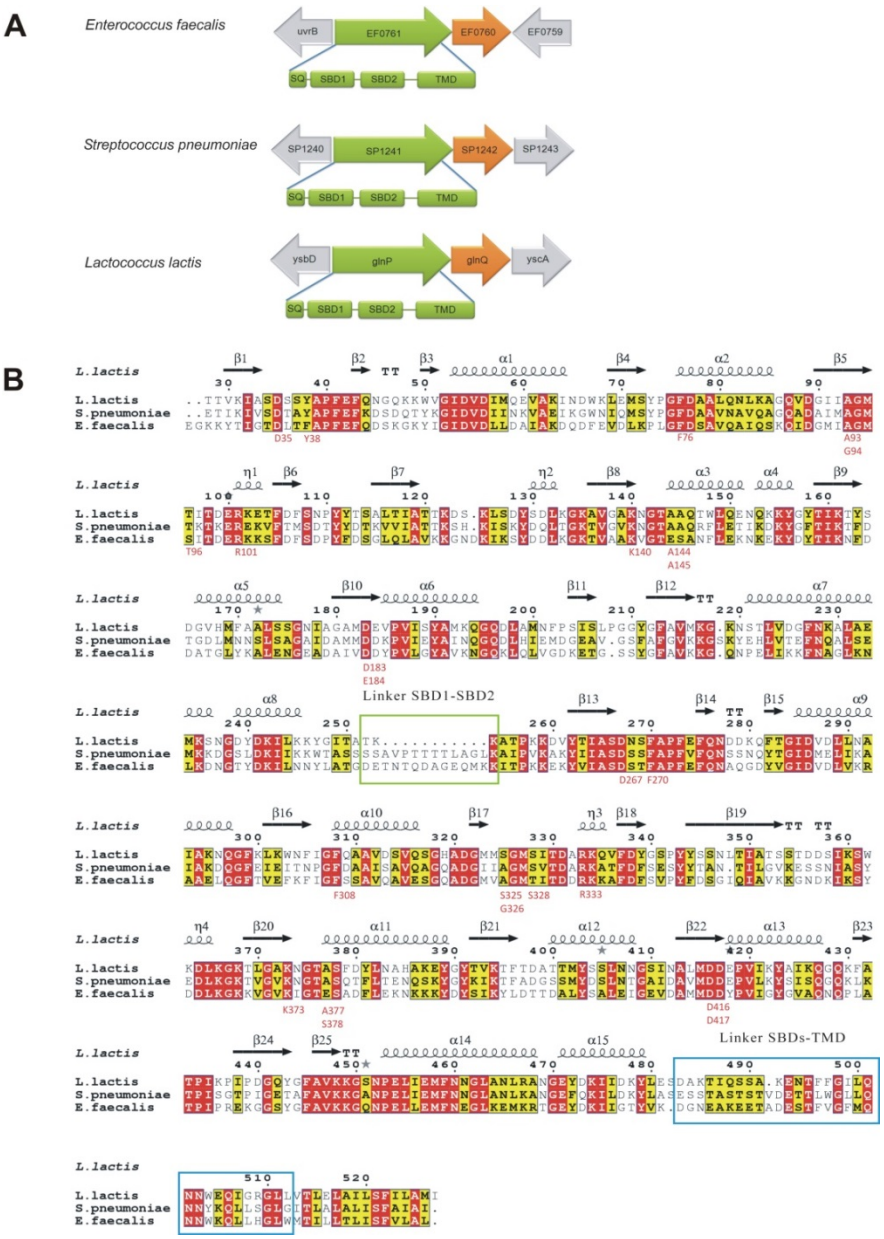

Figure 1. Gene organization of GlnPQ and sequence comparison of SBD of GlnPQ protein. (A) Gene organization of *glnP* (green) and *glnQ* (orange) of the ABC transporters from *Enterococcus faecalis* (EF0761, EF0760); *Streptococcus pneumoniae* (SP1241, SP1242) and *Lactococcus lactis* (*glnP, glnQ*). The flanking genes are indicated. The domain structure of the GlnP protein is indicated: signal sequence, (SQ); substrate-binding domain, (SBD); transmembrane domain (TMD). The neighboring genes (grey) stand for *uvrB* (Excinuclease ABS subunit B); EF0759 (*SapB* protein); SP1240 (uncharacterized hypothetical protein); SP1243 (glucose 6-phosphate dehydrogenase); *ysbD* (uncharacterized

membrane-associated protein) and *yscA* (Lectin-L-type hypothetical protein). (B) Sequence alignment of the substrate-binding domain of the GlnPQ ABC transporters from *L. lactis* IL1403, *S. pneumoniae* TIGR4 and *E. faecalis* V583. The secondary structure and the residue numbers of the tandem SBD from *Lactococcus lactis* are shown above the sequences, and residues in the active site are highlighted in red below the sequences. The alignment was generated using ESPript software (espript.ibcp.fr/). The linker between the tandem SBDs is marked in a green box, while the linker between the SBDs and TMD is depicted in a blue box.

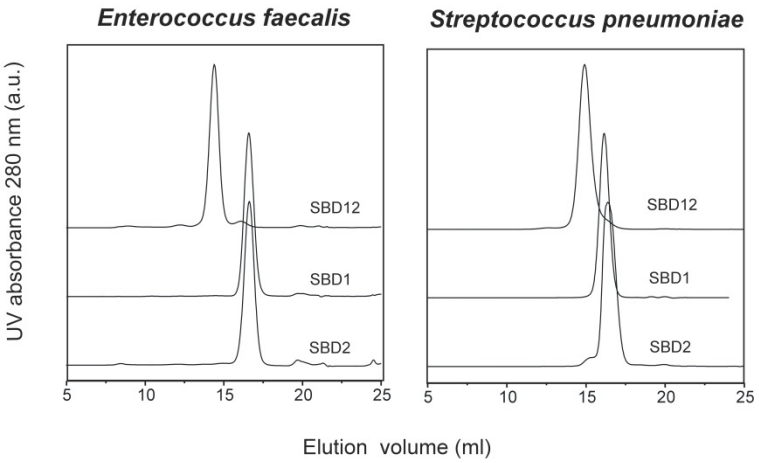

Figure 2. Purified substrate-binding domains. Size-exclusion chromatography (GE Healthcare Superdex 200 10/300 GL column) of purified SBD12, SBD1 and SBD2 from *E. faecalis* and *S. pneumoniae*. The elution profiles of the tandem SBDs (SBD12) and separate domains are consistent with their calculated molecular masses of ~ 53 kDa and ~27 kDa, respectively.

1.2.2 Substrate affinity and specificity

We used isothermal titration calorimetry (ITC) and intrinsic protein fluorescence to characterize the binding of glutamine to the SBD1, SBD2 and SBD12 of GlnPQ from *E. faecalis* and *S. pneumoniae*. ITC is suitable for determining dissociation constants (K_D) in the nanomolar to micromolar range but not for very low affinity binding (submillimolar or higher). In that case fluorescence measurements can be an alternative, provided the intrinsic protein fluorescence changes upon ligand binding. Contrary to ITC, fluorescence measurements do not provide thermodynamic details (ΔH and ΔS) of the binding process. SBD1 of GlnP from *S. pneumoniae* ($SBD1_{Spn}$) has two Trp residues that we used to track conformational changes and to determine the K_D of low affinity binding.

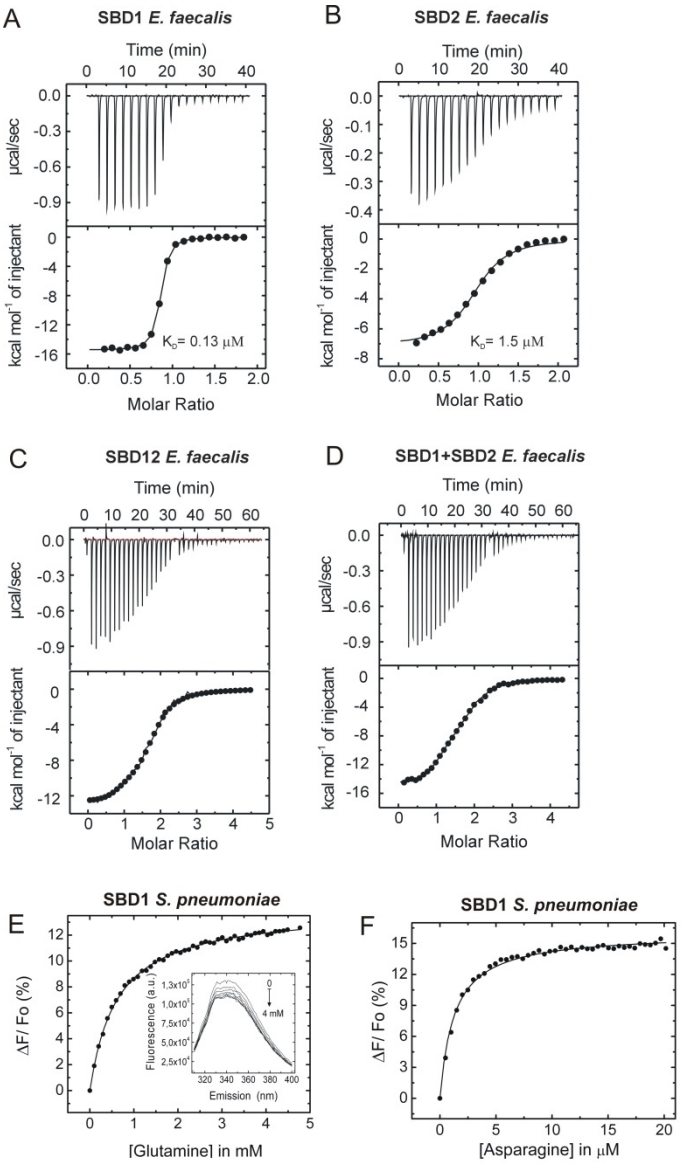

Figure 3. Amino acid binding to SBDs of GlnPQ transporters from *E. faecalis* and *S. pneumonia.* Glutamine binding to SBD1 (A), SBD2 (B), SBD12 (C) and a one-to-one mixture of SBD1 with SBD2 (D) of GlnPQ from *E. faecalis*. Isothermal titration calorimetry (ITC) measurements were performed in 20 mM Na-Mes pH 5.5 plus 150 mM NaCl at 298 K. Purified protein at a final concentration of 35-45 µM was used in the experiments. The upper graphs show the heat released by the protein upon glutamine binding, and the area under each injection signal was integrated and plotted in the lower panel. The solid lines in the lower panel represent nonlinear least-squares fits of the reaction heat for the injection. The enthalpy per mole of glutamine injected is plotted as a function of the protein to glutamine ratio. In case of SBD1 plus SBD2 half the concentration in weight (rather than the sum of the concentration of

SBD1 plus SBD2) was used to allow for comparison with the SBD12 data. Glutamine and asparagine binding to SBD1 of GlnPQ from *S. pneumonia* is shown in panels E and F, respectively. The normalized fluorescence changes at 340 nm were plotted as a function of glutamine or asparagine concentration, yielding K_D values of 716 μM and 1.4 μM, respectively. The emission scans of SBD1 in the absence of glutamine (uppermost trace) and in the presence of successively higher concentrations of glutamine (up to 4 mM) are shown in the inset. The excitation and emission wavelength were 295 and 340 nm, respectively and with a slit width of 1 and 3 nm. The protein concentration was 1 μM and measurements were done in 100 mM Na-Mes pH 5.5 plus 150 mM NaCl at 298 K.

The ITC data for glutamine binding to SBD1, SBD2 and SBD12 from *E. faecalis* are presented in Figure 3 (panels A-C), and the data are summarized in Table 1. $SBD1_{Ef}$ binds glutamine with very high affinity (K_D = 0.13 μM); the K_D for glutamine binding to $SBD2_{Ef}$ is 1.5 μM. The binding isotherms of $SBD12_{Ef}$ are a composite of those of $SBD1_{Ef}$ and $SBD2_{Ef}$, but the corresponding K_D values are difficult to resolve. Importantly the binding isotherm of $SBD12_{Ef}$ is identical to that of a 1-to-1 mixture of $SBD1_{Ef}$ with $SBD2_{Ef}$ (panel C and D), indicating that binding of glutamine to one SBD does not influence binding to the other SBD.

The K_D of glutamine binding to SBD1 from *S. pneumoniae* was too high for ITC measurements and was determined by fluorescence titration instead. Glutamine binding led to the decrease of the protein fluorescence, which was used to estimate the K_D of 716 μM (Figure 3E); this value is three orders of magnitude higher than the K_D of 0.7 μM of the corresponding SBD2 (Table 1). We used the same method to test all natural 20 amino acids as potential substrate (Figure 4). Surprisingly, $SBD1_{Spn}$ binds aspartate (K_D= 410 μM) provided the pH is low (pH 3.9 in our measurements), which suggests that aspartic acid rather than aspartate anion is bound. Another remarkable finding is the high-affinity binding of asparagine (K_D= 1.42 μM; Figure 3F) as compared to glutamine (K_D= 716 μM), while the corresponding SBD2 showed low affinity for asparagine and high affinity for glutamine (data summarized in Table 1). The affinity of $SBD2_{Spn}$ towards asparagine was so low that we could not determine the exact K_D value. The K_D values of glutamine binding to SBD1 and SBD2 of GlnPQ from *L. lactis* are 90 and 0.9 μM, respectively (Table 1). $SBD1_{Ll}$ has high affinity for asparagine (K_D = 0.2 μM), whereas we did not detect any interaction between $SBD2_{Ll}$ and asparagine. In summary, we find that the tandem SBDs of GlnPQ transporters from pathogenic and non-pathogenic Gram-positive bacteria have evolved differently in terms of substrate specificity and affinity. These functional differences occur between GlnPQ transporters from different species but also between the SBDs of a given polypeptide.

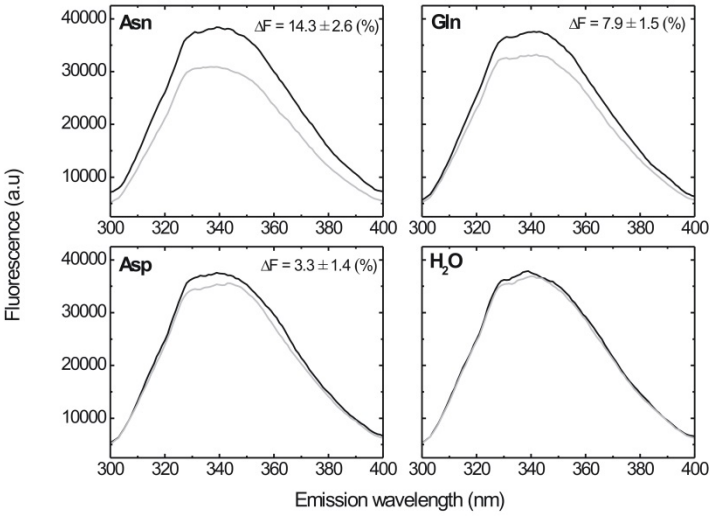

Figure 4. Amino acid binding to SBD1 of GlnPQ from *S. pneumonia*. Protein fluorescence emission scans of SBD1 from *S. pneumoniae* in the absence or in the presence of amino acids (Asparagine, Glutamine or Aspartate) are shown in black and grey lines, respectively. The other 17 amino acids tested gave insignificant fluorescence changes (data not shown). Water (milliQ) was used to correct the fluorescence signal for the dilution effect. Protein fluorescence quenching after each amino acid addition was indicated as ΔF (%). The excitation and emission wavelength were 295 and 340 nm, respectively and with a slit width of 1 and 3 nm. The protein concentration was 1 μM and measurements were done in 100 mM Na-Mes pH 5.5 plus 150 mM NaCl at 298 K.

Table 1. Binding affinity of the SBDs of GlnPQ.

Organism (Ligand)	**K_D (µM)**	**ΔH (kJ/mol)**	**T.ΔS (kJ/mol)**	**ΔG (kJ/mol)**
E. faecalis (Gln)				
SBD1	0.13 ± 0.01	-67.0 ± 2.5	-28 ± 2.7	-39.1 ± 3.7
SBD2	1.49 ± 0.02	-29.5 ± 0.3	3.8 ± 0.4	-33.3 ± 0.5
SBD1 triple mutant	7.13 ± 0.3	-54.7 ± 1	-25.4 ± 0.9	-29.3 ± 1.3
S. pneumoniae (Gln)				
SBD1	716 ± 126#			
SBD2	0.7 ± 0.01	-43.9 ± 0.8	-8.8 ± 0.7	-35.1 ± 1.0
SBD1 triple mutant	0.16 ± 0.03	-26 ± 0.9	12.8 ± 0.8	-38.8 ± 1.2
S. pneumoniae (Asn)				
SBD1	1.42 ± 0.1#			
SBD2	nd			
L. lactis (Gln)				
SBD1	92 ± 16	-23.9 ± 6.3	- 0.9 ± 0.1	-23.0 ± 7.1
SBD2	0.9 ± 0.2	-19.3 ± 2.2	15.4 ± 0.7	-34.7 ± 2.4
L.lactis (Asn)				
SBD1	0.2 ± 0.0	-73.6 ± 1.2	-34.9 ± 1.1	-38.2 ± 1.6
SBD2	no affinity			

Dissociation constants (K_D) and thermodynamic parameters (ΔG, ΔH and T.ΔS) of amino acid binding to SBD1 and SBD2 from *E. faecalis*, *S.* pneumonia and *L. lactis.* nd: not determined due to very low affinity. # based on intrinsic fluorescence measurements.

1.2.3 Crystallization and structure determination

1.2.3.1 Overall structures

We solved crystal structures of SBD1 from *L. lactis* and *E. faecalis* and SBD2 from *L. lactis* (Summarized in Table 2).

Table 2. Data collection and refinement statistics (values in parentheses are for the highest resolution shell)

	E. faecalis	***L. lactis***	***L. lactis***	***L. lactis***	***L. lactis***
	SBD1-liganded	SBD1 –MES	SBD1 -unliganded	SBD2-unliganded	SBD2-liganded
Data Collection					
Space group	$P2_1$	P1	P1	C2	$P2_1$
Cell dimensions					
a, b, c (Å)	40.73, 61.14, 45.12	35.10, 55.63, 55.74	35.09 55.5 56.01	88.69, 89.12, 59.48	42.99, 51.69, 44.23
α, β, γ (°)	90.00, 99.05, 90.00	93.99, 89.71, 97.98	93.35, 92.92, 97.49	90.00, 95.58, 90.00	90.00, 91.10, 90.00
Resolution range (Å)	44.56-1.5 (1.6-1.5)	17.9-1.4 (1.48-1.4)	34.8-1.4 (1.5-1.4)	19.89-1.5 (1.58-1.5)	44.22-0.95 (1-0.95)
Completeness (%)	97.6 (95.6)	93.6 (92.9)	89.1 (87.5)	98.9 (98.2)	98.5 (95.8)
R_{meas} (%)	4.2 (10.4)	3.0 (29.2)	5.6 (34.0)	4.1 (48.9)	4.7 (45.6)
I/ σ(I)	23.8 (14.1)	13.4 (2.5)	6.2 (2.0)	14.3 (1.7)	19.8 (3.2)
Redundancy	3.3 (3.2)	2.0 (2.0)	1.7 (1.7)	2.3 (2.3)	4.2 (3.5)
Refinement					
Resolution range (Å)	36-1.5	17.9-1.4	34.8-1.4	19.89-1.5	44.22-0.95
Number of reflections	34211	76731	91549	72554	119814
R_{work}/R_{free} (%)	12.69/15.50	13.95/18.78	15.57/19.45	14.34/18.59	11.47/11.97
No.atoms					

Protein	1982	3520	6907	3737	4009
Ligand (substrate/buffer/ PEG/ions)	10/12/-/-	-/24/174/-	-/-/49/5	-/-/79/6	20/-/-/-
Water	373	526	627	775	507
Average B-factors (Å^2)					
Protein	12.9	24.6	23.8	16.0	10.7
Water	25.9	39.3	34.3	27.8	15.2
Ligand (substrate/buffer/ PEG)	4.9 /22.1/-/-	-/31.0/48.8/-	-/-/44.6/38.3	-/-/38.2/20.1	4.1/-/-/-
R.m.s deviations					
Bond length (Å)	0.008	0.009	0.009	0.009	0.007
Bond angles (°)	1.205	1.245	1.235	1.289	1.281

The structure of SBD1 of GlnP from *L. lactis* ($SBD1_{Ll}$) was solved in the open-unliganded state at 1.4 Å resolution, whereas structures of $SBD2_{Ll}$ were obtained both in the closed-liganded and open-unliganded states at 0.9 Å and 1.5 Å resolution, respectively. The structure of SBD1 (244 residues) from *E. faecalis* was solved in a closed-liganded conformation at 1.5 Å resolution (Figure 5A). The three SBDs consist of two α/β subdomains: a large one ($SBD1_{Ll}$ residues 29-113, 207-251; $SBD2_{Ll}$ residues 255-345, 441-480 and $SBD1_{Ef}$ residues 1-106 and 207-244) and a small one ($SBD1_{Ll}$ residues 114-206; $SBD2_{Ll}$ residues 346-440 and $SBD1_{Ef}$ residue 111-200). The two domains are connected by two antiparallel β-strands, a common feature of Class II substrate-binding proteins [17]. In the closed-liganded structures the substrate (glutamine) binds in a deep cleft formed by two domains, while the connecting β-strands form the base of the cleft. $SBD1_{Ef}$ has an overall structure similar to $SBD2_{Ll}$ with rmsd of 1.0 Å (Figure 5A).

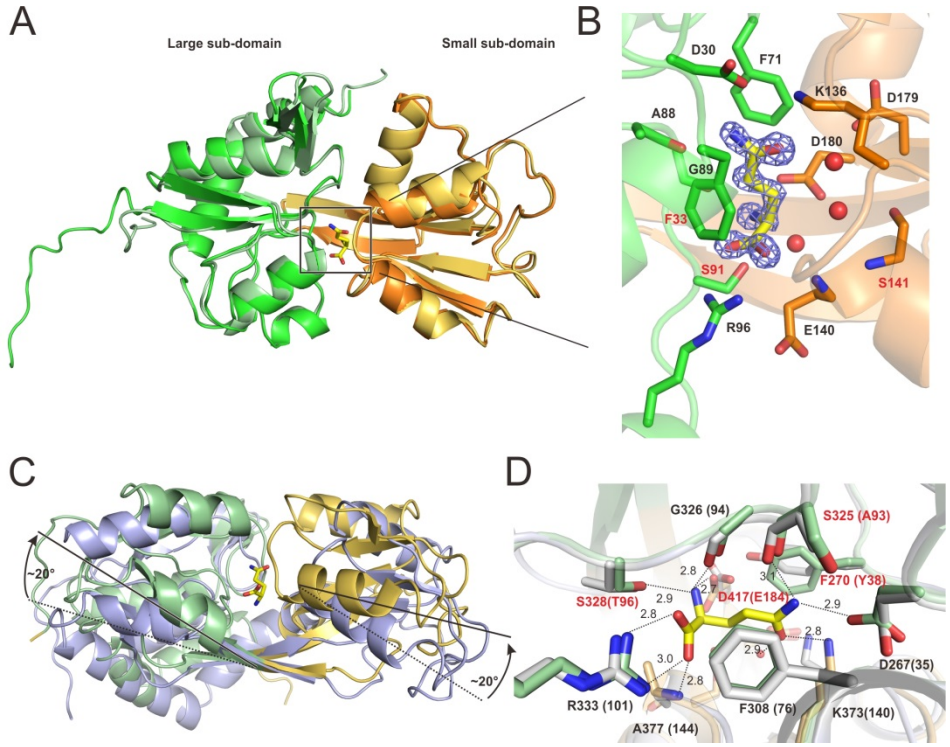

Figure 5. Crystal structures of SBD1 and SBD2 of GlnPQ from *L. lactis* and *E. faecalis*.(A) Superimposition of crystal structures of closed-liganded SBD1 of GlnPQ from *E. faecalis* solved at 1.5 Å and SBD2 of GlnPQ from *L. lactis* solved at 0.9 Å (rmsd ~1Å). The large subdomain is colored in green ($SBD1_{Ef}$) or pale green ($SBD2_{Ll}$) and the small sub-domain is in orange ($SBD1_{Ef}$) or pale orange ($SBD2_{Ll}$). The ligand (glutamine) is shown by stick representation in yellow. (B) Interactions of bound glutamine with the residues in the binding site of $SBD1_{Ef}$; the coloring of subdomain residues is the same as in panel A. Interacting residues are shown as sticks and labeled. Red labels are for the key residues, which change

SBD's affinity toward glutamine. Bound glutamine is colored in yellow and the omit F_o-F_c electron density map, contoured at 2.0 σ level, is shown as blue mesh. (C) Superimposition of unliganded (purple) and liganded (pale green and orange) conformation in $SBD2_{Ll}$. The movements of domains relatively to the hinge region are indicated with black arrows. (D) Comparison of binding sites of liganded-$SBD2_{Ll}$ (green) and unliganded-$SBD1_{Ll}$ (grey). Non-identical amino acids are highlighted in red (Numbering in parentheses is for SBD1). Distances are given in Å.

According to the recent structural classification of substrate-binding proteins [15], the GlnPQ SBDs belong to Subcluster F-IV, which consists of Class II amino acid-binding proteins. DALI searches, using the $SBD2_{Ll}$ structure as a query against the Protein Data Bank, reveals overall highest structural similarity to the *E. coli* glutamine-binding protein GlnBP (PDB code 1WDN), the *Salmonella typhimurium* lysine/arginine/ornithine-binding protein LAO (PDB code 1LAF), the *Geobacillus Stearothermophilus* lysine binding protein ArtJ (PDB code 2PVU) and the *E. coli* histidine-binding protein HisJ (PDB code 1HSL) with DALI Z-scores of ~ 30 for circa 220 aligned Cα atoms. The superposition of $SBD1_{Ef}$ with $SBD1_{Ll}$, $SBD2_{Ll}$, GlnBP, LAO, HisJ and ArtJ indeed shows an overall similar fold with rmsd values between 1.0- 1.6 Å, except for unliganded $SBD1_{Ll}$ that gave a rmsd of 3.1 Å (Figure 6).

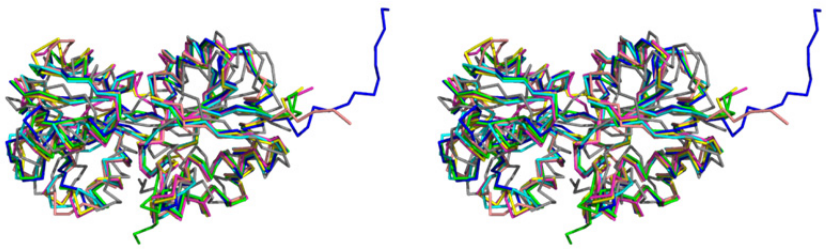

Figure 6. Overall fold of substrate binding proteins from subcluster F-IV. Stereo-view of Cα-backbone superposition of $SBD1_{Ef}$ (blue) with several other SBDs from sub-cluster F-IV; which are $SBD1_{Ll}$ (rmsd 3.12 Å; grey), $SBD2_{Ll}$ (rmsd 1 Å; green), HisJ (rmsd 1.6 Å; magenta), ArtJ (rmsd 1.4 Å; Salmon), LAO (rmsd 1.4 Å; yellow), GlnBP (rmsd 1.3 Å; cyan). The SBDs share a similar fold, a large and a small lobe connected by a linker region

1.2.3.2 Comparison of the open and closed conformations

The open-unliganded and closed-liganded structures of $SBD2_{Ll}$ show that large structural changes occur upon substrate binding. Both domains move as rigid-

bodies about 20° relative to the hinge region made of two β-strands connecting subdomains (Figure 5C). Superposition of the small- and large-subdomains from the $SBD2_{Ll}$ open-unliganded and closed-liganded structures does not reveal significant structural changes within the subdomains upon substrate binding (rmsd ~ 0.4 Å for the aligned Cα atoms). Glutamine bound to SBD2 is completely buried between the large- and the small-subdomain. It is predominantly held in place by hydrogen bonds but additionally it makes ionic and hydrophobic interactions, most of which come from the large-subdomain (Figure 5D). Similar interactions were reported for other amino acid-binding SBDs [18–20], see below.

1.2.3.3 Glutamine-binding site

The SBDs from *L. lactis* have dissociation constants for glutamine of 90 μM (SBD1) and 0.9 μM (SBD2), respectively. To gain insight into the determinants for ligand binding, we compared the structures of these proteins with SBD1 from *E. faecalis*, as $SBD1_{Ef}$ has the highest affinity for glutamine (K_D= 0.13 μM, Table 1). In the closed-liganded state of each protein, L-glutamine is completely buried in a pocket formed between the cleft of the two subdomains, wherein the large subdomain provides the majority of interactions with the ligand (Figure 5B). In $SBD1_{EF}$, the carboxyl group of the bound glutamine is stabilized by a salt bridge with R96, as well as by hydrogen bonds to the backbone nitrogen atoms of S91 and E140 from the large and small subdomain, respectively. The α-amino group makes hydrogen bonds with the hydroxyl group of S91, Oδ2 atom of D180 and the backbone carbonyl of G89, all from the large subdomain. The side chain moiety of the bound glutamine is sandwiched in a hydrophobic pocket formed between F33 and F71. The Nε2 atom of the glutamine forms hydrogen bonds with the Oδ2 atom of D30, and the backbone carbonyl of A88, whereas the glutamine Oε1 atom makes direct hydrogen bonds to Nζ of K136. In addition, the Oε1 atom also makes hydrogen bonds via one and two water molecules with Oδ1 of D179 and the hydroxyl group of S141 side chain, respectively.

In the closed-liganded $SBD2_{Ll}$ structure, the binding pocket is the same as in $SBD1_{Ef}$ with the exception that A88 is replaced by S325 and E140 by A377. These residues however provide the same mode of interaction with the glutamine ligand, namely via backbone carbonyl (A88/S325) and backbone nitrogen (E140/A377) atoms. In the $SBD1_{Ef}$ structure, the side chain of E140 is located in the small subdomain, which most likely tightens the closure of the binding pocket by interacting with two residues in the large subdomain, i.e., via a salt bridge with the side chain of R96 and via hydrogen bonding with the hydroxyl group of T93 (Figure 7). These

interactions may contribute to the higher affinity of the $SBD1_{Ef}$ for glutamine as compared to the $SBD2_{Ll}$ protein.

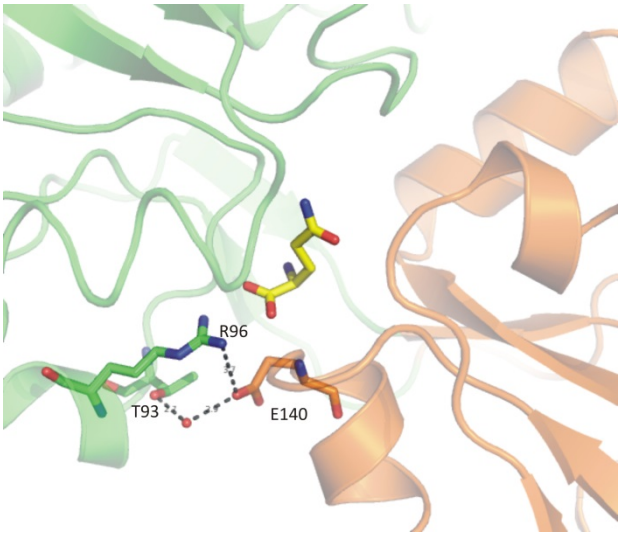

Figure 7. Interaction between the large- and small-subdomain in $SBD1_{Ef.}$ Structural interface between the large (green) and small (orange) subdomains in $SBD1_{Ef.}$; the interacting residues are shown in stick representation. The negatively charged E140 side chain from the small-subdomain makes an electrostatic interaction with the positively charged R96 from the large-subdomain. In addition, the carboxylate side-chain of E140 also makes a hydrogen bond with the hydroxyl moiety of T93 side chain from the large-subdomain. Both R96 and E140 are residues that interact via electrostatic interaction and hydrogen bond with glutamine (colored in yellow) in the binding pocket.

1.2.3.4 Comparison of high and low affinity ligand-binding sites

Similar to the open $SBD2_{Ll}$ structure, one of the $SBD1_{Ll}$ structures is in an open conformation but with a buffer molecule (MES) bound in the position of glutamine. However, binding of MES does not lead to the closure of domains, as the structure is absolutely identical to the one obtained in MES-free conditions. This indicates that interaction of the amino acid side chain is crucial for substrate recognition. To compare the $SBD1_{Ll}$ and $SBD2_{Ll}$ ligand-binding sites, we have superimposed separately the small and large subdomains from SBD1 and the closed glutamine-bound SBD2 (rmsd is 1.0 and 0.9 Å for 127 and 96 aligned Cα atoms, respectively (Figure 5D). The two ligand-binding sites are strikingly similar even though their affinities for glutamine differ by two orders of magnitude. Six out of the nine amino acid residues from $SBD2_{Ll}$ (R333, S328, G326, S325, D267, K373, A377, D416, D417), directly interacting with the ligand, are strictly conserved in $SBD1_{Ll}$. The two non-conserved residues S328 and S325 are replaced by T96 and A93 in $SBD1_{Ll}$. The Oγ1 of T96 in $SBD1_{Ll}$ is fulfilling the same role as the S328 Oγ atom in

$SBD2_{Ll}$. The second interaction is provided by the backbone carbonyl to the Nε2 atom of the bound glutamine, thus this interaction is virtually the same in both structures. The third non-conserved residue in $SBD2_{Ll}$ is D417, which is replaced by E184 in $SBD1_{Ll}$. The two aromatic residues F270 and F308 harboring the side chain of the ligand are conserved in SBD1 as Y38 and F76, respectively. To explain the large differences in affinity of $SBD1_{Ll}$ and $SBD2_{Ll}$ for glutamine, it seems likely that three non-conserved residues play a role in the full closing of the binding pocket, which may not occur in $SBD1_{Ll}$.

1.2.3.5 Comparison with the binding sites of related proteins of subcluster F-IV

The multiple alignments of sequences of $SBD1_{Ef}$, $SBD1_{Ll}$, and $SBD2_{Ll}$ with several other binding proteins within subcluster F-IV (GlnBP, LAO, ArtJ and HisJ; Figure 8) revealed 28-50% identity with $SBD1_{Ef}$. Structural superposition of $SBD1_{Ef}$ and the liganded-form of these proteins show that, out of 19 conserved residues, the equivalent of R96 and D30 form an electrostatic interaction and hydrogen bond with the ligand, respectively, except for HisJ. In HisJ the side chain of the corresponding Asp oriented such that the hydrogen bond cannot be formed. Overall, the binding pockets around the α-carboxylate and the α-amino group of the ligand, and the arrangement of the hydrogen bond framework are well conserved as also discussed for the glutamate/ aspartate binding protein DEBP [21].

On the contrary, the binding pocket for the ligand side chain is poorly conserved, except for the hydrophobic pocket formed by a combination of Phe, Tyr, Trp or Leu residues that sandwich the side chain moiety of the ligand. The residue A88 (in $SBD1_{Ef}$) is substituted by Ser in $SBD2_{Ll}$, HisJ, LAO and ArtJ. The substitution however did not change the interaction mode between the protein and the ligand except in ArtJ. In ArtJ, the hydroxyl group of the corresponding Ser makes water-mediated hydrogen bonds with the Nζ of the lysine ligand (Figure 9A). Replacement of K136 in $SBD1_{Ef}$ with Leu (in HisJ and LAO) abolishes the hydrogen bond interaction with the side chain of the ligand, while in ArtJ this interaction is conserved through the substitution of Lys with Gln. This Lys residue has been suggested to be important for GlnBP specificity [18]. Furthermore, the two residues that provide indirect interaction with the ligand in $SBD1_{Ef}$, S141 and D179, are replaced by Ala ($SBD1_{Ll}$), Gly (ArtJ), Gln (HisJ, LAO) and His (GlnBP), Thr (ArtJ), Gln (HisJ, LAO). The substituted residues provide no interaction with the bound ligand except for the substitution of D179 with His (GlnBP), i.e., a hydrogen bond to the Oε1 atom of the glutamine side chain (Figure 9B). These substitutions may tune the affinity for the respective ligand but it is not possible to predict their roles on

forehand. We thus set out to engineer the specificity of the SBDs, taking advantage of the wealth of structural information described heretofore.

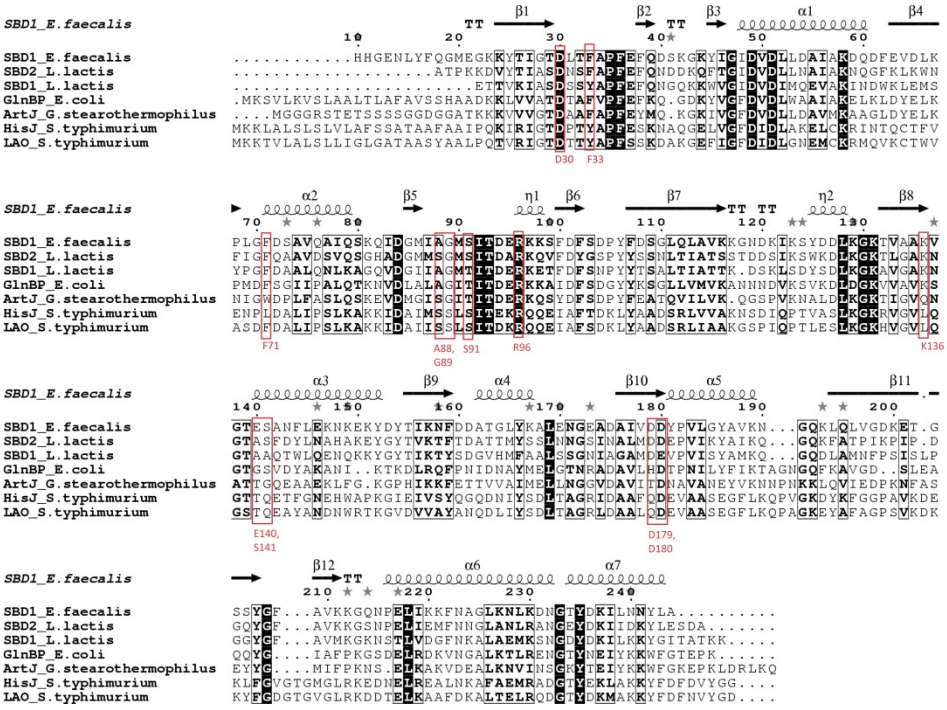

Figure 8. Multiple sequence alignment of $SBD1_{Ef}$ and several other SBDs of known structure from sub-cluster F-IV. Espript was used to generate the secondary structure based on the $SBD1_{Ef}$ crystal structure. The 12 amino acids interacting with the ligand are highlighted and labeled.

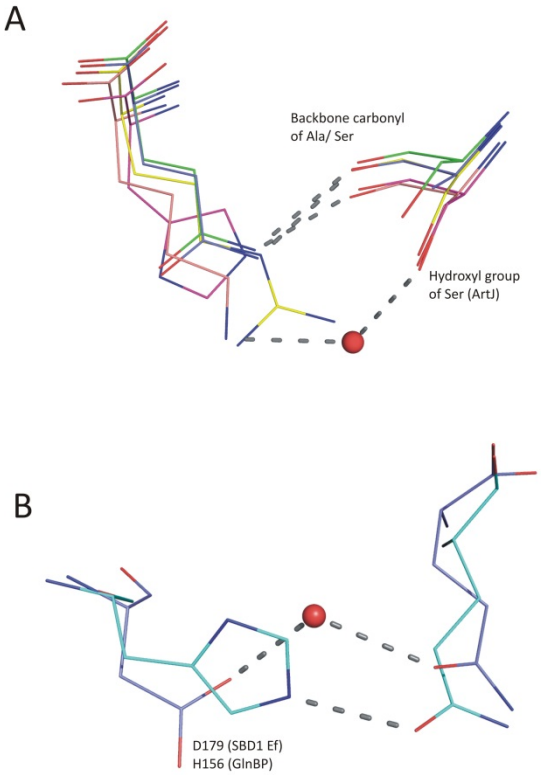

Figure 9. Interaction differences within SBDs of sub-cluster F-IV. (A) The hydrogen-bond interaction between the ligand and the backbone carbonyl of A88 in $SBD1_{EF}$ (blue) is still preserved when the amino acid is replaced by Ser in HisJ (magenta), LAO (yellow), and $SBD2_{LI}$ (green). In ArtJ, the corresponding Ser residue makes a water-mediated hydrogen bond with the ligand via its hydroxyl side chain instead of interacting via a backbone carbonyl as the other proteins do; water molecule is depicted in red sphere. (B) A water-mediated hydrogen bond between the carbonyl group of the glutamine side chain and the carboxylate group of D179 in $SBD1_{Ef}$ (blue) is substituted by a direct hydrogen bond between His and glutamine in GlnBP (cyan).

1.2.4 Engineering of binding affinity

The affinity for glutamine of the SBDs of the GlnPQ ABC transporters from *L. lactis, E. faecalis* and *S. pneumoniae* varies enormously; the lowest and highest K_D differ almost four orders of magnitude (Figure 10). $SBD1_{Ef}$ has the highest affinity and $SBD1_{Spn}$ the lowest affinity. The most notable differences in the binding site residues are highlighted in Figure 10 in orange and green, whereas light blue and blue indicate conserved and highly conserved binding site residues, respectively.

L. lactis		*E. faecalis*		*S. pneumoniae*	
SBD1	**SBD2**	**SBD1**	**SBD2**	**SBD1**	**SBD2**
D35	D267	D30	D273	D28	D271
Y38	F270	F33	F276	Y31	F274
F76	F308	F71	F314	F69	F312
A93	S325	A88	A331	A86	A329
G94	G326	G89	G332	G87	G330
T96	S328	S91	T334	T89	S332
R101	R333	R96	R339	R94	R337
K140	K373	K136	K379	K133	K376
A144	A377	E140	E383	A137	A380
A145	S378	S141	S384	A138	S381
D183	D416	D179	D422	D176	D419
E184	D417	D180	D423	D177	D420
90 µM	0.9 µM	0.13 µM	1.49 µM	716 µM	0.71 µM

Figure 10. The glutamine-binding site of the SBD1 and SBD2 of the GlnPQ transporters from *L. lactis, E. faecalis* and *S. pneumoniae*. The residues indicated are those that interact with the substrate. The most notable differences in the binding site residues are highlighted in orange and green, whereas light blue and blue indicate conserved and highly conserved binding site residues, respectively. The dissociation constants (K_D) for glutamine of each of the SBDs are indicated at the bottom line.

The combined data suggest that F33, S91 and S141 are required for high-affinity binding of glutamine in $SBD1_{Ef}$. In the low affinity $SBD1_{Spn}$, those positions are filled with Y31, T89, and A138. To test this hypothesis, we mutated F33 to Tyr, S91 to Thr and S141 to Ala in the $SBD1_{Ef}$, which should lead to low affinity binding by our prediction. Furthermore, we mutated Y31 to Phe, T89 to Ser and A138 to Ser in the $SBD1_{Spn}$, which should cause the high-affinity binding. Indeed, ITC measurements showed that the K_D for glutamine binding of the $SBD1_{Ef}$ triple mutant was increased from 0.13 to 7.13 µM, whereas the K_D of the $SBD1_{Spn}$ triple mutant was decreased from 720 to 0.16 µM (Figure 11A; Table 1). The high-affinity binding of the $SBD1_{Ef}$ has a large enthalpy term (-67 kJ/mol), which is changed to -54 kJ/mol in the triple mutant, resulting in ~50-fold lower affinity; the entropic term is the same for both proteins. The high-affinity glutamine binding of the $SBD1_{Spn}$ triple mutant has favorable enthalpy and entropy terms, but the ΔH and $T\Delta S$ term of the parental SBD1 could not be determined. The $SBD1_{Spn}$ triple mutant still binds asparagine with relatively high affinity.

The observation that the affinity of the $SBD1_{Spn}$ for glutamine can be increased by four orders of magnitude by introducing only three conservative substitutions is remarkable. This prompted us to investigate the contribution of the individual amino acid substitutions. To perform that, we used the tryptophan fluorescence measurement technique, as the K_D values were too high for ITC. Figure 11B shows that the single mutants have dissociations constants for glutamine in the range

from 40 to 150 μM, which is about an order of magnitude smaller than that of the wildtype $SBD1_{Spn}$ and two orders of magnitude higher than that of the triple mutant. This result demonstrates that each of the residues contributes directly to the binding affinity.

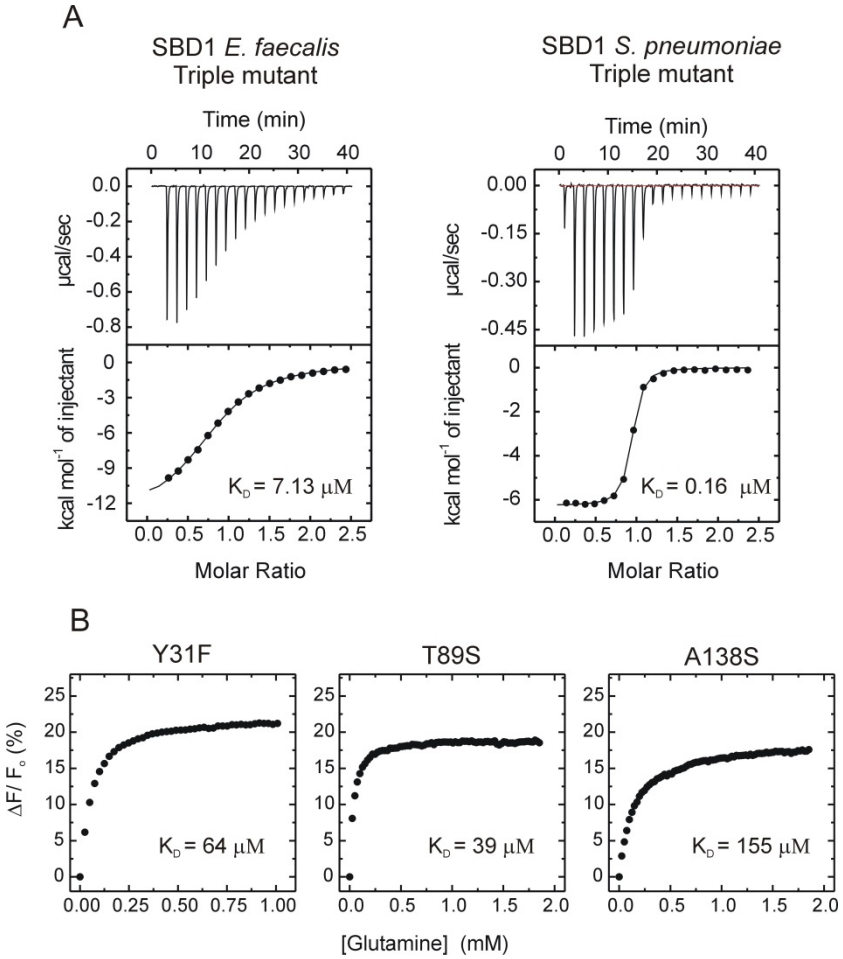

Figure 11. Glutamine binding to the engineered SBDs. (A) Glutamine binding to the *E. faecalis* and *S. pneumoniae* SBD1 triple mutants as determined by isothermal titration calorimetry. The experimental conditions were the same as described in the legend to Figure 3. (B) Glutamine binding to the *S. pneumoniae* SBD1 single mutants (mutations are denoted above the subpanels) as determined by intrinsic protein fluorescence measurements.

1.3 DISCUSSION

To date, the crystal structures of 25 homologues of the *L. lactis, E. faecalis* and *S. pneumoniae* GlnPQ SBDs have been deposited in the protein database bank; for some of these proteins biochemical data on substrate binding are available (summarized in Appendix). However, the SBPs of known structures are either periplasmic (Gram-negative bacteria) or lipid-tethered proteins, in contrast to the aforementioned transporters where two SBDs are fused in tandem and linked to the TMD. Importantly, despite the wealth of structural data and observations that SBPs bind their ligands with nanomolar to sub-millimolar affinity, little is known of the actual affinity and specificity determinants. In addition to the ABC transporter-linked SBDs, a number of crystal structures are available of homologous receptors, which are associated with ionotropic glutamate receptors [22,23]. We now show that ABC transporter-linked SBDs with almost identical overall folds and similar binding pockets can bind glutamine with affinities that different more than four orders of magnitude. Furthermore, we reveal a new feature of the ABC transporters, that is dual substrate specificity owing to the tandem linked SBDs. The *S. pneumoniae* and *L. lactis* GlnPQ use one of the SBDs to bind glutamine with high affinity and asparagine with low affinity and the other SBD having a reciprocal specificity, allowing different amino acids to be accumulated via one and the same transporter. Some SBDs also bind glutamic acid or aspartic acid but not the anionic form of these amino acids, which requires a low pH as they will otherwise be outcompeted by glutamine and/or asparagine.

In ABC transporters with soluble, periplasmic substrate-binding proteins, the transmembrane domain usually interacts with a single receptor [7] but there are exceptions. In the histidine transporter from *Salmonella typhimurium,* the transmembrane domain interacts with two different substrate-binding proteins (HisJ and ArgT), allowing a greater diversity of substrates to be taken [24]. This is analogous to the here described ABC transporters with tandem linked domains. Because the affinity of Type I ABC importers for liganded SBDs is low (about 50 μM; [25–27], the SBD concentration at the site of translocation matters. By linking the SBDs to the TMD, the transporter ensures a high local concentration and thus efficient transport. In the GlnPQ systems, the efficient transport is combined with a broader substrate specificity and a combination of high and low affinity sites within one and the same system, *e.g*. SBD1 from *S. pneumoniae* showed a preference for asparagine (K_D= 1.4 μM) compared to glutamine (K_D= 716 μM), while the corresponding SBD2 showed the opposite selectivity (Table 1). Similarly, SBD1 from *L. lactis* binds asparagine with high affinity (K_D= 0.2 μM) and glutamine with low affinity (K_D= 92 μM), whereas SBD2 only accepts glutamine (and glutamic

acid). It is possible that the high affinity site allows the scavenging of amino acids when nutrient conditions are limited, whereas the low affinity site allows for faster transport when amino acids are available in excess [28]. In this way the GlnPQ transporters combine several properties that allow the uptake of different amino acids that are needed at high concentrations in the cell, under conditions of varying nutrient availability.

Biological activity of a molecule is driven by free energy changes with enthalpic (static interactions) and entropic (dynamic interactions) contributions. The ITC data on the binding of glutamine to wild type and mutant $SBD1_{Ef}$ indicate that the negative ΔH term is the determinant for high-affinity ligand binding. Compared to the $SBD2_{Ef}$, the $SBD1_{Ef}$ has a more favorable (more negative) ΔH and less favorable (more positive) ΔS. Our structural analysis shows that the binding sites for amino acids in the different SBD1 and SBD2 proteins are similar and from the structures alone it is not evident what determines the specificity and/or ligand affinity. It is possible that low affinity ligands are unable to fully close the binding site and thus form fewer interactions with the protein than high-affinity ligands. In case of glutamine, the extra methylene group could render the ligand too bulky thus preventing full closure of the binding pocket of $SBD1_{Ll}$, which would in turn result in a lower affinity. In case of $SBD2_{Ll}$, the closed binding pocket may be larger allowing more interactions between the protein and glutamine than between the protein and asparagine. Here, one could speculate that too few interactions between asparagine and the protein preclude full closure of the binding pocket. In fact, we have preliminary single-molecule FRET data that point to a correlation between the extent of SBD closure and high affinity binding.

Protein-ligand interactions are determined by hydrogen bonding and electrostatic, electron-π, hydrophobic and van der Waals interactions between a protein and its substrate. In case of $SBD1_{Ef}$, eight direct hydrogen bonds between side chain/backbone residues and glutamine are present. In addition, salt bridges and electron-π interactions contribute to the binding. However, it is also known that conformational entropy, i.e. internal dynamics, contributes to the protein-ligand interactions and thus to protein activity and binding affinity [29]. Slow internal motions, related to poorly populated conformational states, can affect activity in a manner that is not readily predicted from static X-ray structures. Our mutational analysis of high- and low-affinity SBDs suggests that differences in the binding of glutamine are largely due to differences in amino acids in the binding pocket rather than second-shell residues or long-range effects. Remarkably, the amino acid substitutions that lead to large changes in the binding affinity are all conservative, and the effects are not readily predicted from the sequences. A

similar property has been described for the glycine betaine/proline binding protein from *Archaeoglubus fulgidus*; here, a single mutation in the binding site enhanced binding affinity by four orders of magnitude [30].

For *L. lactis*, a non-pathogenic Gram-positive bacterium related to *E. faecalis* and *S. pneumoniae*, the GlnPQ transporter is essential for growth in amino acid-containing media. *L. lactis* as well as many Gram-positive pathogens require glutamine or glutamate for biosynthesis and osmoregulation. Accumulated glutamine is readily converted into glutamate and the latter is generally the most abundant amino acid in the bacterial cytoplasm [31]. All these requirements warrant an efficient transport of glutamine and glutamate. In *B streptococci,* GlnPQ plays a role in virulence by affecting the regulation of expression of fibronectin adhesion[32]. Deletion of the glutamine transporters SPD 1098/1099 in *S. pneumoniae* D39 (equivalent gene to SP1241/1242 in *S. pneumoniae* TIGR4) diminished bacterial fitness and virulence [33]. Moreover, mutations on *glnP* from *S. pneumoniae* significantly reduced the adhesion ability of the bacterium to human pharyngeal epithelial cells, suggesting an important role of the GlnPQ ABC transporter in host colonization [34]. The natural auxotrophy for Gln/Glu and the requirement for GlnPQ to take up these amino acids makes this transporter a promising target for drug/antibiotic development. Finding an effector of the SBDs that abolishes transport would make the compound effective without the need to enter the cell. Currently, we are exploring the structures of the GlnPQ domains to design and test small molecule inhibitors of glutamine and glutamate transport.

In conclusion: we show that tandem SBDs allow single ABC transporters to capture different amino acids with high and low affinity, making the process flexible and efficient. We have engineered the specificity determinants of the ABC transporters and show that a few conservative substitutions in the active site of the SBDs can increase the substrate affinity by 3-4 orders of magnitude. Because the binding sites have a very similar architecture, the molecular determinants of high-affinity binding are most probably valid for SBDs of other ABC transporters from subcluster F-IV. The point mutations in the same architecture of binding sites, coupled with small, yet sufficiently different ligand size, suggest control of SBD closure and provide an elegant way to utilize one and the same transporter for uptake of several amino acids. The ability to discriminate glutamine (and glutamic acid) from asparagine by nearly the same SBDs allows the efficient transport without direct competition. It is likely that dual specificity and affinity is a general feature of other subfamilies of ABC transporters, for which our study may serve as starting point for further investigations.

1.4 MATERIALS AND METHODS

1.4.1 Construction of plasmids

The genes coding for the GlnPQ-derived substrate binding domains (SBDs) were amplified by PCR from genomic DNA of *Enterococcus faecalis* V583, *Streptococcus pneumoniae* TIGR4 and *Lactococcus lactis* NZ9000 and nLIC complementary primers [35]. The first 23-28 amino acids corresponding to the signal sequence were omitted. The PCR products were digested with *Swa*I, followed by treatment with T4 DNA polymerase and then inserted into the *Swa*I and T4 DNA polymerase-treated pBADnLIC vector (containing the sequence coding for the N-terminal His_{10} tag), using ligation-independent cloning. The resulting plasmids pBadnLIC-SBDs (the subscripts Ef, Spn and Ll refer to SBDs from *E. faecalis, S. pneumoniae* and *L. lactis*, respectively) were transformed to *E. coli* MC1061. The parental pBADnLIC-$SBD1_{Spn}$ was subsequently used to generate the following substitution mutants: Y31F, T89S and A138S, using the Quikchange PCR protocol (Stratagene). All the cloned fragments were verified by DNA sequencing (ServiceXS, The Netherlands). The triple mutants (Y31F, T89S, and A138S) of $SBD1_{Spn}$ and $SBD1_{Ef}$ (F33Y, S91T, S141A) were obtained by chemical synthesis of the corresponding genes (Invitrogen).

1.4.2 Overexpression of GlnPQ-derived substrate binding domains

E. coli MC1061, expressing wild type or mutant variants of GlnPQ-SBDs, was cultivated aerobically in Luria Broth medium supplemented with 1% (w/v) glucose plus 10 µg/ml ampicillin at 37°C. The cells were induced using $2 \cdot 10^{-3}$ % (w/v) L-arabinose at $OD_{600} \sim 0.5$ and growth was continued for 2 more hours. Cells were harvested by centrifugation at 10,000xg for 30 min at 4°C, washed once and resuspended in 50 mM potassium phosphate (KPi) pH 7.0 plus 150 mM NaCl (buffer A) and stored at -80°C.

1.4.3 Protein purification

The frozen cells were thawed at room temperature, and diluted to an OD600 ~ 100 with buffer A. Subsequently, 100 µg/ml deoxyribonuclease type I, 10 mM $MgSO_4$ plus 0.1 mM phenylmethanesulfonyl fluoride (PMSF) were added. The cells were broken by a single pass in a cell disruptor at 25,000 psi and 5°C (Constant system Ltd). The cell lysate was mixed with 5 mM Na-EDTA to prevent protein degradation. The unbroken cells and cell debris were removed by ultracentrifugation 267,000xg for 90 min at 4°C. The supernatant was collected and 0.5 ml Nickel-Sepharose resin (Amersham Biosciences) was added per 10 ml

cell lysate. The mixture was incubated for 1 h at 4°C in 50 mM KPi pH 8, 200 mM KCl plus 10 % (v/v) glycerol (buffer B) supplemented with 20 mM imidazole. Subsequently, the resin was poured into a 10 ml disposable column (BioRad) and washed with 20 column volumes (CV) of buffer B containing 50 mM imidazole. The protein was eluted in three fractions of one CV, using buffer B supplemented with 500 mM imidazole. The second elution fraction (containing most of the purified protein) was loaded onto a Superdex 200 gel filtration column (GE Healthcare) equilibrated with buffer C (20 mM Na-Mes pH 5.5 and 150 mM NaCl). Fractions containing SBD protein were collected and concentrated to ~ 20 mg/ml, using Vivaspin 30 or 50 kDa molecular weight cut off (MWCO). For binding experiments, an additional step was employed before eluting the protein. The potentially bound ligands were removed by partial unfolding of the protein while bound to the Ni-Sepharose resin, followed by refolding [36,37]. The following wash steps were performed consecutively to partially unfold-refold the protein: 40 CV of 2 M Guanidine-HCl (GndHCl), 4 CV of 1.5 M GndHCl, 4 CV of 1 M GndHCl, 4 CV of 0.5 M GndHCl and finally 8 CV of 0 M GndHCl (all in buffer B as basal solvent).

For SBDs of GlnPQ from *L. lactis* the His-tag was cleaved off by His-tagged-TEV protease at a ratio of 1:40 (w/w) with respect to the purified protein, subsequently the protein was dialyzed against 50 mM Tris-HCl pH 8.0, 0.5 mM EDTA plus 0.5 mM DTT overnight at 4 °C. The Histidine-tagged TEV and residual uncut protein was removed using 0.5 ml bed volume of Ni^{2+}-sepharose. The flow through of the column was concentrated, using a Vivaspin concentrator (GE Healthcare) with a 10 kDa cut off. Concentrated protein was loaded to a Superdex-200 column (GE Healthcare) and eluted with 20 mM Hepes-NaOH, pH 7.5 plus 150 mM NaCl. Crystallization trials were set up immediately after purification, whereas for other purposes the proteins were stored at -80 °C after flash freezing in liquid N_2.

1.4.4 Isothermal titration calorimetry

ITC experiments were conducted using the ultrasensitive ITC200 calorimeter (MicroCal) at 25°C. Glutamine in buffer C (approximately 40 µl) at a concentration of 350-500 µM was added stepwise into the temperature-equilibrated ITC cell filled with ~200 µl of SBD protein in the same buffer and at a concentration of 35-50 µM. The experiments were repeated at least three times. Control measurements included titration of buffer C into the protein solution.

Data were analyzed by using the non–linear curve fitting functions for one binding site [38], provided by the ORIGIN-based software of MicroCal. The calculated curve was obtained from the best fitting parameters and was used to determine the molar enthalpy change for protein-ligand complex formation, stoichiometry (n),

and the corresponding association constant (K_A). The dissociation constant (K_D) is defined as $1/K_A$, and the standard free energy change of binding, ΔG = - RT ln(K_A). The molar entropy change, ΔS, was calculated from $\Delta G=\Delta H-T\Delta S$.

1.4.5 Fluorescence spectroscopy

Substrate binding was measured on Spex Fluorolog 322 fluorescence spectrophotometer (Jobin Yvon) in a stirred quartz cuvette at 25°C. Purified SBD1 from *S. pneumoniae* was diluted in 100 mM Mes pH 5.5 plus 150 ml NaCl to a concentration of 1 µM (final volume 1 ml) and incubated for 5 min under mild stirring to reach the equilibration temperature before stepwise addition of the substrate (or buffer as control). In case of aspartate binding experiments, the buffer was changed to 50 mM Na-Acetate pH 3.9 plus 150 mM NaCl. The substrate was added in 1 µl steps, using a pump (Harvard apparatus) fitted with a 500 µl gastight glass syringe (Hamilton Co.). The syringe was connected to the cuvette by tubing with an internal diameter of 0.13 mm (Vici AG International). The excitation wavelength was 295 nm and the emission was measured at 340 nm with a slit width of 1 and 3 nm, respectively. The signal was measured for 20 sec following 5 sec of mixing time to assure full equilibration. Fluorescence titrations were analyzed as described previously [28], and curve fitting was performed in ORIGIN software.

1.4.6 Crystallization and structure determination of SBD1 and SBD2 from *L. lactis*

All crystals were grown using the hanging drop vapor diffusion method at 281 K. Drops were prepared by mixing protein and the reservoir solutions in a 1:1 v/v ratio. Crystals of $SBD1_{Ll}$ (concentrated to 18.0 mg/ml) grew in 0.1 M MES pH 6.0, 5% PEG3000 and 30, 35 or 40% PEG400 within 5 days. To obtain MES-free crystals, reservoir solution was exchanged to 40% PEG 600, NaH_2PO_4/Citric Acid, pH 4.2. The reservoir solution from which crystals of liganded-SBD2 (concentrated to 22.5 mg/ml) were grown contained 0.1 M CH_3COONa, pH 5.0, 30% PEG6000, 100 mM $CaCl_2$ plus 2 mM glutamine. The $SBD2_{Ll}$ crystals grew in 1-2 months. Crystals were cryo-protected in a solution similar to the reservoir solution but supplemented with 40% PEG6000. Unliganded-$SBD2_{Ll}$ crystals were grown from reservoir solution containing 0.1 M CH_3COONa pH 5.0, 200 mM $MgCl_2$ and 40% PEG6000 within 1-3 days.

The $SBD1_{Ll}$ datasets were collected at beam line ID23-1, ESRF, Grenoble, France, while the liganded and unliganded-$SBD2_{Ll}$ datasets were collected at beam line ID14-1 of the same facility. All datasets were processed using the XDS package and

SCALA [39–41]. The $SBD1_{LI}$ crystals belong to space group *P*1 with two molecules per unit cell and 38% solvent content. The liganded- and unliganded-$SBD2_{LI}$ crystals belong to space groups $P2_1$ (one molecule per asymmetric unit and 36% solvent content) and *C*2 (two molecules per asymmetric unit and 46% solvent content), respectively.

All structures were solved by molecular replacement method, using the program Phaser 2.1.4 of the CCP4 program suite [42]. To solve the structure of the unliganded-$SBD2_{LI}$, the structure of the glutamine-binding protein GlnH from *E. coli* was used as a search model (PDB code: 1GGG). The model was divided into two subdomains: subdomain 1 (residues 5-88 and 182-244) and subdomain 2 (residues 89-181). The program Phaser found the rotation and translation solutions for three out of four subdomains present in the asymmetric unit (two subdomains 1 and one subdomain 2). Inspection of the crystal packing and the electron density map indicated the location of the second subdomain 2. To find solutions for the second subdomain 2, the spherically-averaged phased translation function, phased rotation function and phased translation function options of the program MOLREP from the CCP4 program suite was used [41,43]. Further model building and corrections were done using the program COOT [44]. The models were refined using Phenix [45] with 5% of reflection randomly set aside to monitor the refinement progress. The overall quality of the model was assessed using the program MolProbity [46]. Final refinement statistics are shown in Table 2. The crystal structures of SBD2 in closed Gln bound and SBD2 in open unliganded conformations have been deposited to the RCSB PDB with the PDB IDs 4KQP and 4KR5, respectively. SBD1 structures have been deposited under the PDB ID 4KPT (with bound MES) and 4LA9 (MES-free).

1.4.7 Crystallization and structure determination: SBD1 from *E. faecalis*

Initial crystallization trials were set up as vapour–diffusion sitting drops at 278 K, using various commercially available screens from Molecular Dimension (JCSG+, Structure Screen, Cryo Screen, and Pact Premier) and the Mosquito crystallization robot (TTP LabTech) for drop dispensing. Droplets were composed of 0.1 µl protein stock solution (20 mg/ml) in buffer 20 mM Na-Hepes pH 7.5 and 200 mM NaCl and 0.1 µl crystallization solutions, and the droplets were equilibrated against 100 µl of the crystallization solution. Initial crystals of SBD1 from *E. faecalis* were small needles or sea urchin–shaped crystals and obtained at several conditions. Optimization of the crystallization conditions gradually improved the crystal physical properties and finally yielded a single needle or plate crystals diffracting to 1.4 Å. The crystals appeared within a week and grew to full size in 2 weeks.

Crystals ranging from 50-500 µm, could be grown at 278 K from solutions containing 100 mM MMT buffer system [47] pH 4-6 and 26-28 % PEG 1500. Crystals were soaked in mother liquor supplemented with 33 % PEG 1500 and flash-frozen in liquid nitrogen. For co-crystallization with glutamine, the ligand (100 mM stock in milliQ water) was mixed with the protein solution yielding a final concentration of 5 mM.

The data were collected at the beamline PX I/ X06SA, Swiss Light Source, Switzerland. Data processing and reduction were carried out using MOSFLM [48], XDS [40] and scaled using SCALA from CCP4 package [49]. The structure of $SBD1_{Ef}$ with glutamine bound was solved by molecular replacement using Phaser [42], and using the previously determined structure of the closed-liganded $SBD2_{Ll}$ as search model. The solutions from Phaser were subjected to automated model building, using ARP/wARP [50], and completed by applying several cycles of manual model building, using COOT [44]. Subsequently, the output from COOT was further refined, using Refmac5 [51] and Phenix [45]. The crystal structure of SBD1 from *Enterococcus faecalis* in closed Gln bound has been deposited to the RCSB PDB with the PDB ID 4G4P.

ACKNOWLEDGEMENTS

This work was supported by grants from the Netherlands organization for scientific research (NWO, Top-subsidy grant 700.56.302 to BP, Vici grant to DJS) and the EU (EDICT program). Authors are grateful to beam line personnel at 14-1, ID 23-1 (ESRF, Grenoble) and at X06SA (SLS, Villigen) for technical assistance.

REFERENCES

1. Rees, D. C., Johnson, E. & Lewinson, O. ABC transporters: the power to change. *Nat. Rev. Mol. Cell Biol.* **10,** 218–227 (2009).
2. Erkens, G. B., Majsnerowska, M., ter Beek, J. & Slotboom, D. J. Energy Coupling Factor-Type ABC Transporters for Vitamin Uptake in Prokaryotes. *Biochemistry (Mosc.)* **51,** 4390–4396 (2012).
3. Rodionov, D. A. *et al.* A novel class of modular transporters for vitamins in prokaryotes. *J. Bacteriol.* **191,** 42–51 (2009).
4. Erkens, G. B. *et al.* The structural basis of modularity in ECF-type ABC transporters. *Nat. Struct. Mol. Biol.* **18,** 755–760 (2011).
5. Orelle, C., Ayvaz, T., Everly, R. M., Klug, C. S. & Davidson, A. L. Both maltose-binding protein and ATP are required for nucleotide-binding domain closure in the intact maltose ABC transporter. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 12837–12842 (2008).
6. Oldham, M. L. & Chen, J. Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* **332,** 1202–1205 (2011).
7. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev. MMBR* **72,** 317–364, table of contents (2008).
8. Böhm, S., Licht, A., Wuttge, S., Schneider, E. & Bordignon, E. Conformational plasticity of the type I maltose ABC importer. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 5492–5497 (2013).
9. Bao, H. & Duong, F. ATP Alone Triggers the Outward Facing Conformation of the Maltose ATP-binding Cassette Transporter. *J. Biol. Chem.* **288,** 3439–3448 (2013).
10. Korkhov, V. M., Mireku, S. A. & Locher, K. P. Structure of AMP-PNP-bound vitamin B12 transporter BtuCD-F. *Nature* **490,** 367–372 (2012).
11. Rice, A. J. *et al.* EPR spectroscopy of MolB2C2-a reveals mechanism of transport for a bacterial type II molybdate importer. *J. Biol. Chem.* **288,** 21228–21235 (2013).
12. Joseph, B., Jeschke, G., Goetz, B. A., Locher, K. P. & Bordignon, E. Transmembrane Gate Movements in the Type II ATP-binding Cassette (ABC) Importer BtuCD-F during Nucleotide Cycle. *J. Biol. Chem.* **286,** 41008–41017 (2011).
13. Berger, E. A. & Heppel, L. A. Different Mechanisms of Energy Coupling for the Shock-sensitive and Shock-resistant Amino Acid Permeases of *Escherichia coli*. *J. Biol. Chem.* **249,** 7747–7755 (1974).
14. Van der Heide, T. & Poolman, B. ABC transporters: one, two or four extracytoplasmic substrate-binding sites? *EMBO Rep.* **3,** 938–943 (2002).

15. Berntsson, R. P.-A., Smits, S. H. J., Schmitt, L., Slotboom, D.-J. & Poolman, B. A structural classification of substrate-binding proteins. *FEBS Lett.* **584,** 2606–2617 (2010).

16. Geertsma, E. R., Groeneveld, M., Slotboom, D.-J. & Poolman, B. Quality control of overexpressed membrane proteins. *Proc. Natl. Acad. Sci.* **105,** 5722–5727 (2008).

17. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286,** 279–290 (1999).

18. Sun, Y.-J., Rose, J., Wang, B.-C. & Hsiao, C.-D. The structure of glutamine-binding protein complexed with glutamine at 1.94 Å resolution: comparisons with other amino acid binding proteins1. *J. Mol. Biol.* **278,** 219–229 (1998).

19. Oh, B. H. *et al.* Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J. Biol. Chem.* **268,** 11348–11355 (1993).

20. Trakhanov, S. *et al.* Ligand-free and -bound structures of the binding protein (LivJ) of the Escherichia coli ABC leucine/isoleucine/valine transport system: trajectory and dynamics of the interdomain rotation and ligand specificity. *Biochemistry (Mosc.)* **44,** 6597–6608 (2005).

21. Hu, Y. *et al.* Crystal structure of a glutamate/aspartate binding protein complexed with a glutamate molecule: structural basis of ligand specificity at atomic resolution. *J. Mol. Biol.* **382,** 99–111 (2008).

22. O'Hara, P. J. *et al.* The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins. *Neuron* **11,** 41–52 (1993).

23. Jin, R. *et al.* Crystal structure and association behaviour of the GluR2 amino-terminal domain. *EMBO J.* **28,** 1812–1823 (2009).

24. Higgins, C. F. & Ames, G. F. Two periplasmic transport proteins which interact with a common membrane receptor show extensive homology: complete nucleotide sequences. *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6038–6042 (1981).

25. Prossnitz, E., Gee, A. & Ames, G. F. Reconstitution of the histidine periplasmic transport system in membrane vesicles. Energy coupling and interaction between the binding protein and the membrane complex. *J. Biol. Chem.* **264,** 5006–5014 (1989).

26. Dean, D. A., Hor, L. I., Shuman, H. A. & Nikaido, H. Interaction between maltose-binding protein and the membrane-associated maltose transporter complex in *Escherichia coli. Mol. Microbiol.* **6,** 2033–2040 (1992).

27. Doeven, M. K., Abele, R., Tampé, R. & Poolman, B. The binding specificity of OppA determines the selectivity of the oligopeptide ATP-binding cassette transporter. *J. Biol. Chem.* **279,** 32301–32307 (2004).

28. Lanfermeijer, F. C., Picon, A., Konings, W. N. & Poolman, B. Kinetics and consequences of binding of nona- and dodecapeptides to the oligopeptide binding protein (OppA) of *Lactococcus lactis*. *Biochemistry (Mosc.)* **38,** 14440–14450 (1999).

29. Tzeng, S.-R. & Kalodimos, C. G. Protein activity regulation by conformational entropy. *Nature* **488,** 236–240 (2012).

30. Tschapek, B. *et al.* Arg149 is involved in switching the low affinity, open state of the binding protein AfProX into its high affinity, closed state. *J. Mol. Biol.* **411,** 36–52 (2011).

31. Poolman, B., Smid, E. J., Veldkamp, H. & Konings, W. N. Bioenergetic consequences of lactose starvation for continuously cultured *Streptococcus cremoris*. *J. Bacteriol.* **169,** 1460–1468 (1987).

32. Tamura, G. S., Nittayajarn, A. & Schoentag, D. L. A Glutamine Transport Gene, glnQ, Is Required for Fibronectin Adherence and Virulence of Group B Streptococci. *Infect. Immun.* **70,** 2877–2885 (2002).

33. Härtel, T. *et al.* Impact of glutamine transporters on pneumococcal fitness under infection-related conditions. *Infect. Immun.* **79,** 44–58 (2011).

34. Kloosterman, T. G. *et al.* Regulation of Glutamine and Glutamate Metabolism by GlnR and GlnA in *Streptococcus pneumoniae*. *J. Biol. Chem.* **281,** 25097–25109 (2006).

35. Geertsma, E. R. & Poolman, B. High-throughput cloning and expression in recalcitrant bacteria. *Nat. Methods* **4,** 705–707 (2007).

36. Vahedi-Faridi, A. *et al.* Crystal structures and mutational analysis of the arginine-, lysine-, histidine-binding protein ArtJ from *Geobacillus stearothermophilus*. Implications for interactions of ArtJ with its cognate ATP-binding cassette transporter, Art(MP)2. *J. Mol. Biol.* **375,** 448–459 (2008).

37. Staiano, M. *et al.* Unfolding and refolding of the glutamine-binding protein from *Escherichia coli* and its complex with glutamine induced by guanidine hydrochloride. *Biochemistry (Mosc.)* **44,** 5625–5633 (2005).

38. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179,** 131–137 (1989).

39. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763 (1994).

40. Kabsch, W. *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 (2010).

41. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242 (2011).

42. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

43. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 22–25 (2010).

44. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501 (2010).

45. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 (2010).

46. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21 (2010).

47. Newman, J. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 610–612 (2004).

48. Leslie, A. G. W. & Powell, H. R. in *Evolving Methods for Macromolecular Crystallography* (eds. Read, R. J. & Sussman, J. L.) 41–51 (Springer Netherlands, 2007). at <http://link.springer.com/chapter/10.1007/978-1-4020-6316-9_4>

49. Leslie, A. Recent changes to the MOSFLM package for processing film and image plate data. *Oint CCP4 ESF-EAMCB Newsl. Protein Crystallogr.* **26,** (1992).

50. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3,** 1171–1179 (2008).

51. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 355–367 (2011).

Appendix

Overview of amino acid-binding proteins available in the protein data bank (PDB).

#	Protein/ gene name	Organism	Ligands	Conformation		Highest resolution	PDB code	Highest affinity
				unliganded	liganded			
1	*glnH*	*Escherichia* coli	Gln	v	v	1.94	1GGG, 1WDN	0.5 µM
2	*glnH*	Burkholderia *pseudomallei*	Gln		v	2.4	4F3P	nd
3	ArtJ	*Geobacillus stearothermophilus*	Arg, Lys, His		v	1.79	2Q2A,2Q2C, 2PVU	0.039 µM
4	stm4351	*Salmonella* enterica	Arg		v	1.9	2Y7I	2.9 µM
5	GluR0	*Nostoc punctiforme*	L-Glu		v	2.1	2PYY A, 2PYY B	25 µM
6	wS0279	*Wolinella succinogeneshelp*	Lys		v	2.62	3K4U	-
7	LAO	*Salmonella enterica*	Lys, Arg, Ornithine	v	v	1.8	1LAF, 1LST, 1LAH, 1LST, 2LAO, 1LAG	14 nM
8	HisJ	*Escherichia coli*	His		v	1.89	1HSL A, 1HPB	40 nM
9	ArtJ	*Legionella pneumophila*	Arg	v		2.06	3KGZ	-
10	stu0877	*Streptococcus thermophilus*	Polar amino acid			1.9	3HV1	-
11	ct381	*Chlamydia trachomatis*	Arg			1.92	3DEL	-
12	cbu_0482	*Coxiella burnetii*	Arg		v	1.59	3TQL	-

13	cpn0482	*Chlamydia pneumoniae*	Arg		v	2.1	3N26	-
	cpb0502	*Chlamydia pneumoniae*	Arg		v	2	3QAX	-
14	yckb_bacsu	*Bacillus subtilis*	-	v		2.2	2IEE	-
15	spo1306	*Silicibacter Pomeroyi*	His/Glu/Gln/Arg/Opine		v	2	3I6V	-
16	GluR0	*Synechocystis sp.*	Ser		v	1.9	1IIT	240 nM
			Glu		v	1.6	1II5, 1IIW	
17	DEBP	*Shigella flexneri*	Glu/ Asp		v	1	2VHA	1.6 µM
18	DEBP	*Shigella flexneri.*	-		v	1.5	2IA4 A,B	-
19	DalS	*Salmonella enterica*	D-Ala		v	1.9	4DZ1	-
20	stm1633	*Salmonella enterica*	D-Ala		v	2.14	4F3S	-
21	ngo 0372	*Neisseria gonorrhoeae*	Cys	v	v	2.26	2YJP	21 nM
22	cjaA	*Campylobacter jejuni*	Cys		v	2	1XT8	100 nM
23	PEB	*Campylobacter jejuni*	Asp		v	1.49	2V25	1.9 µM
24	Spd_0109	*S. pneumoniae Canada Mdr19A*	Arg, Glutathione		V	1.05	4I62, 4H5F, 4EQ9	-
25	lmo239	*Listeria monocytogenes*	CYS		V	1.45	4GVO	-

SBD1 of GlnP from E. faecalis was used as a query for the protein blast search.
* no publication available, only crystal structure deposited in PDB.